MAYANK GOLEY

Los Angeles | mayankgoley055@gmail.com | (714) 519-7734
linkedin.com/in/mayankgoley | github.com/mayankgoley | mayankgoley.me

SUMMARY

Software developer and AI engineer with 3+ years building data intensive microservices at Capgemini (energy/utilities) and ML-serving infrastructure at Infosys (CPG/demand forecasting). Core stack: Python, Java, Spring Boot, FastAPI, Kafka, PostgreSQL, Kubernetes. M.S. Computer Science at Cal State Fullerton (Dec 2026).

TECHNICAL SKILLS

Languages: Python, Java, JavaScript, TypeScript, SQL
Backend & Frameworks: FastAPI, Django, Flask, Spring Boot, Spring Batch, REST APIs, SOAP
AI & ML: PyTorch, scikit-learn, NumPy, Pandas, LangGraph, Qdrant, RAG
Data & Messaging: PostgreSQL (full-text search/tsvector, PostGIS), Oracle Database, Redis, Kafka
Cloud & DevOps: AWS (ECS, S3, SageMaker, CloudWatch), Docker, Kubernetes, CI/CD
Frontend: React, Next.js 14, HTML/CSS
Testing & Payments: pytest | Stripe Connect

PROFESSIONAL EXPERIENCE

Infosys | Machine Learning Engineer — Nov 2022 – Jun 2023
Client: PepsiCo — Demand Forecasting Platform | Chandigarh, India

- Created PySpark data transformation modules across 8 regional SAP configurations, handling locale specific date formats, currency conversions, tax structures, and SKU hierarchy mappings with full test coverage
- Wrote SQL pipelines joining SAP sales (VBAK/VBAP), inventory (MARD), and material (MARC) tables with promotional calendars, converting base units to business aggregates across SKU location combinations
- Implemented an outlier detection and classification system using IQR flagging with rule-based routing against trade calendars, integrating a planner review queue back into the ML training pipeline

Capgemini | Software Developer — Feb 2020 – Jul 2022
Client: Hydro One — Grid Analytics & Energy Platform | Pune, India

- Developed Java batch pipelines (Spring Batch) ingesting 2M+ daily smart meter readings via SFTP, with timestamp validation and anomaly detection, reducing processing failures by 40%
- Optimized Spring Boot REST endpoints with JPA query tuning and response caching, cutting API response time by 35% for the customer-facing consumption and outage portal
- Maintained cross system integration pipelines between GIS, OMS, CIS, and MDM using SOAP services and SFTP, resolving 95%+ of data sync failures within SLA
- Rewrote compliance reporting SQL with indexing and date-range partitioning, reducing NERC/PUC report generation from 2 hours to 40 minutes

PROJECTS

TeachWise | Flask, Next.js 14, PostgreSQL, Redis, Stripe Connect, Jitsi — teachwiseedu.com

- Engineered a full stack tutoring marketplace with 4 role based dashboards, 59 routes, and 35 PostgreSQL tables handling bookings, wallet transactions, and guardian approval chains for minor students
- Integrated Stripe Connect for split payments with automated tutor payouts, PostGIS geo aware tutor search (tsvector full-text), and Redis backed search history, on a Next.js 14 App Router frontend with TypeScript strict mode and auth middleware

Financial Intelligence System | LangGraph, Qdrant, FastAPI, React — github.com/mayankgoley/financial-intelligence-system

- Built a hybrid RAG pipeline combining Qdrant dense vector search with BM25 keyword search via Reciprocal Rank Fusion, with fiscal year metadata filtering for SEC filing analysis
- Designed a multi agent debate architecture (LangGraph) with Bull, Bear, Synthesis, and Verifier agents, with conditional retry when the Verifier detects hallucinated claims

Traffic Sign Classification | PyTorch, scikit-learn, GTSRB — github.com/mayankgoley/ML_projects

- Trained a scratch CNN reaching 96.9% test accuracy on the 43 category GTSRB German traffic sign dataset, beating a frozen ImageNet pretrained ResNet18 at 66.7%, with inverse frequency class weights handling the 10x class imbalance
- Identified low resolution domain shift as the failure mode where ImageNet features did not transfer to small low detail traffic signs, with speed limit signs (20, 30, 50, 70, 80 km/h) consistently the weakest classes

Energy Forecasting | PyTorch, NumPy, Pandas — github.com/mayankgoley/ML_projects

- Built a 24 hour ahead energy load forecasting model on 16 years of hourly PJM regional grid data (145,000 rows) with chronological train/val/test splits, train only scaler fitting, and look ahead leakage assertions
- Compared LSTM and Transformer encoder architectures against a one week naive baseline, achieving 58% MAE reduction (3519 to 1466 MW, 4.6% MAPE on 32000 MW mean load), with Transformer beating LSTM by only 37 MW MAE (within noise)

EDUCATION

California State University, Fullerton | M.S., Computer Science | GPA: 3.6 — Aug 2024 – Dec 2026
Dr. A.P.J. Abdul Kalam Technical University | B.Tech., Information Technology — Aug 2015 – Jun 2019

CERTIFICATIONS

Oracle Cloud Infrastructure 2025 Certified Generative AI Professional (ID: 323311678OCI25GAIOCP), 2025
Oracle Cloud Infrastructure 2025 Certified Developer Professional (ID: 103017713OCID25CP), 2025